Word-final vowel shortenings in the Rigveda and the nominative singular of PIE *eh2- and *ih2-stems

Litscher, Roland

Abstract: The common reconstruction of *eh2- and devī-type *ih2-stem singular nominatives as *-ah2 (< *-eh2) and *-ih2 includes no overt marking of the nonneuter nominative singular. Instead of taking this for granted and searching for a motivation in the history of those stems, we raise the possibility of a reconstruction with overt marking (*-āh2, *-īh2). In the second part, the distribution of some word-final vowel shortenings in the Rigveda is analysed in order to determine whether they correlate with the different origins of the vowels or not. The distributions prove interesting, and call for further investigation, but since shortening seems possible irrespective of the origin of the vowel, the Vedic evidence cannot help decide between the two possible reconstructions.

DOI: https://doi.org/10.1515/if-2015-0015

Roland Litscher

Word-final vowel shortenings in the Rigveda and the nominative singular of PIE *eh_2- and *ih_2-stems

Abstract: The common reconstruction of *eh_2- and *devī́*-type *ih_2-stem singular nominatives as *$-ah_2$ (< *$-eh_2$) and *$-ih_2$ includes no overt marking of the non-neuter nominative singular. Instead of taking this for granted and searching for a motivation in the history of those stems, we raise the possibility of a reconstruction with overt marking (*$-āh_2$, *$-īh_2$). In the second part, the distribution of some word-final vowel shortenings in the Rigveda is analysed in order to determine whether they correlate with the different origins of the vowels or not. The distributions prove interesting, and call for further investigation, but since shortening seems possible irrespective of the origin of the vowel, the Vedic evidence cannot help decide between the two possible reconstructions.

Keywords: *eh_2-stems, *ih_2-stems, *devī́*-stems, nominative singular, lengthened grade, feminine gender, word-final vowel shortening, Rigveda

Roland Litscher: University of Zurich; litscher@indoger.uzh.ch

1 Reconstruction of *eh_2- and *ih_2-stem nominatives

1.1 Introduction

Most directly reconstructed singular nominatives of masculine or feminine gender in late PIE are usually posited with either a segmental morpheme *$-s$, a lengthened vowel in the word-final morpheme, or both. The further reconstruction of the origins and the exact distribution of the allomorphs is still a matter of dispute, but most of these issues are of only minor relevance to the topic discussed in the present paper. Crucial are not the origins of the allomorphy, but rather its mere existence: textbook examples like nominative *$ph_2tḗr$ vs. vocative *$ph_2tér$ clearly show that the lengthened suffix vowel is an overt marker of nominative case, singular number, and non-neuter gender in the inflectional system of late PIE much in the same sense as its segmental allomorph *$-s$. Of course, the alterna-

10.1515/if-2015-0015

tion with *-*s* and the morphological contrast with zero-marked short-vowel forms (voc. sg. m./f. and nom./acc. sg. n.) do not necessarily imply that both must historically be mere phonological variants of each other,[1] or that all cases with the combined markers *long vowel* + *-*s* must be "reconstructed away" to yield a perfect complementary distribution of the markers.[2] What is implied, however, by the overall stability of the opposition to truly zero-marked vocatives and neuters is that the lengthened-grade nominatives are not mere phonological variants of unmarked or zero-marked forms (as is claimed by e. g. Beekes 1985: 151–166), because in that case the absence of lengthening in forms such as **ph₂tér* as opposed to lengthened **ph₂tḗr* would call for an explanation.[3]

While most non-neuter nominatives reconstructed for PIE are thus morphologically marked in some way, there seems to be one major class of exceptions: the nominatives of **eh₂*- and *devḗ*-type **ih₂*-stems are commonly reconstructed as *-*ah₂* (< *-*eh₂*) and *-*ih₂* with no overt marking of the non-neuter nominative singular. This leads to the question to be investigated in the following sections: What are the possible reasons for this systematic gap in the reconstructed distribution of nominative marking? Either this distribution must have had some motivation at some point in the history of PIE morphosyntax – which may or may not be recoverable – or it is based on misleading reconstructions.[4] When the question has been raised at all, scholars so far have sought to answer it along the lines of the former possibility. Section 1.2 is a discussion of some of these approaches that have in my opinion not yet provided a convincing scenario. While this is not in itself a valid argument against the possibility of a historically and functionally motivated distribution, it still suggests that the other option could reward closer consideration. That there is indeed a possible alternative to the prevailing reconstruction is argued for in section 1.3. It turns out that a reconstruction with lengthened vowels – and thus overt marking of the nominative – is in fact as compatible with the reflexes in the individual languages as the conventional short-vowel forms.

Finally, one important set of potential evidence for one of the alternatives is investigated in more detail in section 2. Since the work of Kuiper (1947, 1955) it is

1 Though in my opinion this is still the most probable scenario.

2 To be sure, many forms with combined markers are clearly innovations of individual languages, but there is no reason to believe that all examples are late, or the result of the addition of the segmental marker *-*s* to an already marked long-vowel form.

3 I cannot see anything "attractive" in the "very daring" explanation presented in Beekes 1985: 100f.

4 A third line of reasoning – that we are dealing with an unexplainable morphological quirk – represents the first approach too, just paired with a conviction that the motivation will not be recoverable.

usually assumed that some of the variation in word-final vowel length revealed by the prosodic patterns of the Vedic hymns is historically rooted in sandhi variants of earlier sequences of short vowel plus laryngeal. It therefore seems possible that the reflexes of postulated sequences of long vowel plus laryngeal differed both from those of short vowel plus laryngeal and from those of original long vowel without laryngeal. Although suggestive, the absence of written variation between long and short vowels in nominatives of *ā*- and *ī*-stems (section 2.1) is no conclusive argument against a preform with short vowel, because there are also important groups of instrumentals and neuter plurals without variation. On the other hand, there is also no evidence against a long vowel preform from the distribution of metrically short vowels in prevocalic position in the Rigveda (section 2.2). The vowels of some groups do behave differently in this position, but the original vowel length and the presence of a laryngeal are not conditioning factors. Shortening is attested for all investigated groups and the differences in frequency within historically similar groups are actually bigger than between groups.

1.2 Possible motivations for zero-marked nominatives

The main approach holding the view that the non-overt marking of **eh*$_2$- and **ih*$_2$-stem nominatives are motivated within the morphological system of PIE is that of the Leiden school. The unmarked "nominative" (or in their scenario rather absolutive) is seen as an archaism that remained unaffected by the introduction of the "nominative" (ergative) ending **-s*. In accordance with his version of "hysterodynamic" inflection, Beekes 1985: 20–36 reconstructs **-h*$_2$ (not **-eh*$_2$) as the PIE form. It would thus be an archaism not only in its lack of overt marking, but also in its being an unaccented, zero-grade suffix in the nominative.[5] Obviously the question raised in the present paper is simply moot within the Leiden system. But that of all things the nominative of **h*$_2$-stems must be so archaic seems very improbable to me, since most scholars consider animate **h*$_2$-stems to belong to a rather late stage of PIE.[6] Furthermore, zero-marked animate nominatives may make perfect sense within the elaborate system of Beekes and the Leiden school, but basic assumptions underlying this system must be accepted into the bargain: that (early) PIE once had an ergative alignment system; that lengthened grades de-

5 See the very useful overview of the possible subtypes and their historical offshoots in his system in Beekes 1985: 161.

6 This is, to be sure, not a problem for Beekes himself, since in his view the common ancestor of the neuter plural nominative-accusative and the singular nominative of **h*$_2$-stems was a feminine derivative from the beginning (Beekes 1985: 28).

veloped (exclusively?) from phonetic lengthening before resonants and in monosyllabic forms; and that the "hysterodynamic" (in the Leiden sense of the term, comprising the amphi- or holokinetic and hysterokinetic types of the Erlangen model) and proterodynamic types share a common origin.[7] In my opinion, none of these three assumptions is convincing, but this is not the place to contrast the two models, which both have their merits and weaknesses.[8] The present author's views on the system of PIE nominal morphology share more of the basic assumptions of the Erlangen model, and since the necessity or feasability of the reconstruction under discussion is only relevant within the context of that model, I will concentrate on problems and improvements pertaining thereto.

In a model of PIE nominal inflection that interprets lengthened-grade nominatives as overtly marked forms on a par with those with segmental *-*s*, the lack of nominative marking in *h_2-stems is typically taken to be a remnant of an earlier stage without differentiation of nominative and accusative in the precursor of these forms. There is a considerable amount of disagreement about the earlier categorical status of these forms, and about when and how they were remade into the later feminines,[9] but most scholars would probably agree that the differentiated nominatives and accusatives of those stems are rather late creations within PIE. While the question of the reconstruction of the *h_2-stem nominatives is thus obviously intertwined with the question of the origins of the IE feminine gender, it is important to stress that it arises at some point in most scenarios, and is thus not tied to one particular theory of IE gender. Why were only the accusatives, but not the nominatives provided with an overt ending when these forms were created?

The only proposal that gives a principled reason for this is Tichy (1993). Tichy argues that the first pair of forms with this marking pattern was demonstrative *$sáh_2$ and *$táh_2$-*m* created from *$táh_2$ on the model of **só* and **tó-m*. The main problem with this approach is its consequence that the agreement forms used to mark feminine gender on pronouns and thematic adjectives would have to be older than the forms of animate nouns in *-h_2.[10] Since the latter, but not the former, are already well attested in Anatolian (see now especially Melchert 2014), this explanation seems unsatisfactory.

7 See Litscher 2014: 142–144 for a very different proposal about the formal connections between the amphi- and hysterokinetic types.

8 Hopefully, the 2013 Leiden conference marks the beginning of a renewed discussion of the basic concepts of the competing views, ideally converging on a better reconstruction.

9 See Litscher 2009 and 2014 for the present author's views on these matters and the other chapters of Neri & Schuhmann 2014 for an overview of current discussions.

10 See, inter alia, Fritz 1998: 259–260 and Litscher 2009: 272–274 for criticism of the view that the demonstrative is the starting point for the development of the feminine gender.

All other current scenarios for the creation of feminine gender provide no specific reason for the absence of marking in the nominative as opposed to the accusative.[11] These include also the proposal of Fritz (1998) that is intended as a specification of Tichy's approach, and that of Harðarson (1987: 100–103) that shares many of Tichy's assumptions, although the explanations of the feminine agreement forms are in fact markedly different. Starting from originally neuter singulars lacking a nominative-accusative distinction, a differentiation of nominative and accusative forms must have been introduced at some point. Why was this achieved by marking just the accusative? In other recent approaches to the problem that start from nouns in *-h_2- of already animate gender giving rise to feminine agreement through their reinterpretation as adjectives (Kim 2014, Melchert 2014, Nussbaum 2014: 303–306), the assumption of an unmarked nominative beside a marked accusative also demands an explanation.[12] Why should those nouns unlike all other animate nouns lack a nominative marker?

1.3 An alternative reconstruction

The lack of answers to the questions raised in the preceding section spurred me to look at the issue from a different angle. Maybe the problem is not the lack of answers, but rather that the wrong questions have been asked. Is it possible that these forms were in fact marked for nominative singular? Such an alternative reconstruction is indeed feasible, because the pattern with nominative *-*s* and accusative *-*m* was not the only model available for differentiated nominatives and accusatives during the renewal of these forms. Instead of introducing *-*s* into the nominatives of the evolving animate *h_2-stems, speakers built those forms on the model of the hysterodynamic inflection,[13] where the nominative was marked by the lengthened grade of the suffix vowel. A reconstruction with nominative *-$\bar{a}h_2$[14] and *-$\bar{\imath}h_2$, accusative *-ah_2-*m* and *-ih_2-*m*, and vocative *-ah_2-∅ and

11 Of course, it remains perfectly possible that speakers at the critical time felt the need to mark the accusative but not the nominative, for some unknown reason.

12 Animate *h_2-stems at some stage of the development are also assumed in the otherwise strongly diverging models of Ledo-Lemos 2000, Luraghi 2009, 2014, and Litscher 2009, 2014.

13 The assumption that truly hysterodynamic forms from stems in *-ih_2- existed earlier is unwarranted. Reconstructions such as a nominative *-$i\bar{e}h_2s$ (e. g. Mayrhofer 1986: 133) are needed neither for the Latin nominatives in *-ies* (see Klingenschmitt 1992: 127–135) nor for the *vr̥kī*-stems in general (see Widmer 2005).

14 Given that it is a relatively late creation within PIE, one would expect this form to have been created after the colouring of **eh_2 to *ah_2. The question of the age of laryngeal colouring is so

*-ih_2-∅ is thus well embedded in the system of PIE noun inflection. The advantages of such a reconstruction are first and foremost that most of the scenarios cited in section 1.2 are rendered much more coherent, as the animate *h_2-stems need not be seen as some morphologically aberrant in-between category; rather they were a perfectly normal type of noun, as soon as they were created.

Individualizing uses of (animate) *eh_2-stems have attracted some attention beyond the discussions of the creation of the feminine gender. In an inspiring talk at the 14th Fachtagung of the Indogermanische Gesellschaft at Copenhagen (2012), Hannes Fellner and Laura Grestenberger have convincingly argued that the second members of Latin compounds of the *agricola*-type and similar formations in Greek (e.g. βαθυδίνης 'deep-eddying') and Armenian reflect neither abstract nor collective and neither specifically feminine nor neuter nouns, but rather deadjectival formations with a suffix in *-h_2- of animate gender and in individualizing function. Similar functions of *-ah_2- are also claimed by Melchert (2014) and Nussbaum (2014) for Anatolian. These functions are all better understandable if animate *eh_2-stems were productive prior to and to some extent independent of the emergence of the feminine gender.[15]

Another area where the alternative reconstruction allows for a better understanding of the facts concerns the short-vowel vocatives of stems in *-h_2 attested in several IE dialects.[16] It is of course perfectly possible to explain these forms with Kuiper by the process of laryngeal loss in pausa, combined with the observation that vocatives tend to be treated as extrasentential elements: under these assumptions, the shortened variants from laryngeal loss in pausa were frequently vocatives, but only rarely nominatives, and therefore were redistributed to yield short-vowel vocatives opposed to long-vowel nominatives. Even so, reconstructing an earlier opposition of nominatives in long vowel plus *h_2 to vocatives in short vowel plus *h_2 gives an additional edge to this statistical argument: whereas the vocatives do result in short vowels in pausa and before vowels, the nomina-

intertwined with differing opinions about the existence of PIE *a and the validity of Eichner's Law that I will not delve into that issue here. The proposed development is only problematic if one both considers colouring to be very late and accepts Eichner's Law. This amounts to assuming an independent development in every single branch, which I deem to be the most unlikely of the different possibilities.

15 The scenario proposed in Litscher 2009; 2014 differs from Melchert's only insofar as the individualizing *eh_2-stems are seen not as the ultimate source of the feminine gender, but rather as a by-product of the first steps towards it. The other (crucial) point of disagreement is of course about when and how the agreement of forms in *-h_2 arose.

16 E. g. Ved. voc. *devi* vs. nom. *devī́*, Gr. voc. νύμφᾰ vs. nom. νύμφη, OCS *ženo* vs. nom. *žena*; see Kuiper 1947: 210.

tives had no such variants at all, since the long vowel would have remained even when the laryngeal was lost in pausa or before a following vowel. To arrive at the attested distribution, only the (probably rare) long variant of the vocative had to be abandoned. This seems all the more natural considering that the original opposition of short vocative to long nominative, and the parallelism with hysterodynamic stems, were thus restored.

In most IE languages, there is no reason to expect different reflexes for long vowel plus laryngeal versus short vowel plus laryngeal.[17] One possible exception is Balto-Slavic, where at least according to some views on the history of their vocalism (see Kortlandt 1985, 1997) a difference might be expected.[18] If original lengthened grade vowels are indeed distinguished by their circumflex intonation from the reflexes of short vowels plus laryngeal with acute intonation, then the acute of the Lithuanian nominatives could provide an argument against the reconstructions with morphological long vowels. But since the contrast of circumflex and acute intonation in diphthongs is usually traced back to the absence versus presence of a following laryngeal, a similar effect can be held responsible for the acute in the cases under discussion too: the acute would by these assumptions result from the presence of a laryngeal after the original long vowel.

2 Word final long vowels in Vedic

2.1 Short vowels from vowel plus laryngeal in the written text

In his in-depth investigation of some cases of alternation in vowel length in the transmitted text of the Rigveda, Kuiper (1955) has shown convincingly that they may have arisen from different treatments of word-final laryngeals in different positions. The most transparent cases are perhaps the gerunds in *-ya* and *-tya* (1955: 7–10). Since the padapāṭha has short *-a* everywhere, it is clear that these forms were considered to be underlyingly short not only in the later language,

17 For instance, the assumption of something like **-$\bar{\imath}h_2$* > **-$\underset{\circ}{\bar{\imath}}\underset{\circ}{h_2}$* for Greek and Tocharian does not cause serious additional complications. While such a development is certainly less expected in the case of a long vowel than a short one, it is still perfectly possible, since there are no principled reasons why the reflexes of long and short *i* preceding a vocalic segment should differ.

18 These issues are by no means resolved yet (see Hock 2004: 13–21 for an overview), but Kortlandt's view is taken for granted here for the sake of argument. If the alternative view (that lengthened grades yield acute vowels) should be correct, the Balto-Slavic forms would be no obstacle at all for the reconstruction proposed here.

but already at the time of the padakāra. Besides 19 contracted instances,[19] the saṃhitāpāṭha attests 51 tokens[20] with long and 24 with short vowels in a remarkable distribution (see Table 1).

Table 1. Uncontracted *-(t)ya* in the Rigveda saṃhitā

	total	end of even pada	end of odd pada	before caesura	opening	cadence
-(t)ya	24	15	–	3	3	3
-(t)yā	51	–	13	15	16	7

The high number of long variants in openings (predominantly in third syllables of 11- and 12-syllable padas) and before a caesura (many in fifth syllables of 11- and 12-syllable padas) show that these are not cases of metrical lengthening, but rather original long vowels shortened in some positions. Kuiper explains the short forms by assuming laryngeal loss in pausa, and suggests that the neat distribution at the end of padas is the work of later editors.[21] Originally, we should thus expect short forms at the end of padas and, if not contracted, before vowels too, but long ones elsewhere. For the position before vowels, the evidence is indecisive: the only line with a hiatus demanded by the metre has *ya* as the fifth syllable of a jagatī (10.53.3c *átaḥ saṃgr̥bhyābhibhūta ā́ bhara*). This leaves us with nine exceptional instances of short forms before a consonant within the pada, which are all from the first and tenth book. The cases in openings and in front of the caesura may in principle reflect alterations by later editors, because a long vowel would give more canonical lines, but all of them are not uncommon with a short vowel either. The restriction to Books One and Ten, however, suggests a different interpretation, as a later editor would probably have altered forms in the other books too. That not only the later editors and the padakāra, but also at least some of the poets of Book Ten considered those forms to be underlyingly short is confirmed by

19 This includes the two occurrences of contracted *-tya* not mentioned by Kuiper (10.99.5d *abhī́tya*, and 10.66.14c *étya*), but excludes 1.166.9b (*mithaspr̥dhyeva taviṣā́ṇy ā́hitā* |), which is better taken as a neuter plural. The exceptional long *ā* in the padapāṭha suggests that this was also the analysis of the padakāra.

20 Kuiper's 50 instances plus 1.56.1d (*rátham āvr̥tyā háriyogam r̥bhvasam* ‖), which is accidentally missing in Zubatý 1889: 298–300.

21 Why the long variants of the gerunds have been introduced at the end of odd padas, but not those of the neuter plurals (see below), remains an open question (Kuiper 1955: 9, 18–19).

the three attestations in cadences in Book Ten, where the metre clearly demands short syllables.[22]

Similar variation is also attested for neuter plural nominative-accusatives in *-ī/i* and *-ū/u* (Kuiper 1955: 12–26) and instrumentals in *-(t)ī/i* (Kuiper 1955: 2–5, 11–12). While it seems evident that we are dealing with basically the same phenomenon in these cases, the distribution of the variants in these groups is harder to assess exactly. First, there are many more cases of doubtful interpretation, since many of these forms may also belong to different morphological categories; and second, editorial alterations have been applied in a less clear-cut fashion.

The instrumentals differ from both the gerunds and the neuter plurals in two important ways. Apart from *śámī̆* and *suśámī̆*, all examples are either long or short throughout, and short variants seem to have been replaced by long ones, but not vice versa. Short forms are transmitted almost exclusively at the end of padas,[23] and are restricted to forms without long variants. But whenever a long-vowel variant does exist, this is also used in pada-final position. The probable original distribution is still intact in the case of *śámī* (3×) and *suśámī* (2×) within padas vs. *śámi* (3×) and *suśámi* (1×) at the end.

The neuter plurals differ from the gerunds in that later editorial work introduced the short variants also within the pada on a large scale. This was systematically the case before consonant groups and the caesura.[24] At the time of editorial rehandling these forms were thus obviously believed to be underlyingly short like the gerunds, and the introduction of these variants has gone even further, as short variants are widely attested in second syllables too. In these cases lexical and formulaic considerations are also relevant.[25] It remains unclear, however, whether some of the poets already considered the short neuter plurals to represent the "true" form, because the surprisingly few examples of metrically justified short vowels before consonant are of a rather exceptional character.[26] Also in contrast

22 10.71.9c *tá eté vā́cam abhipádya pāpáyā* (9th/12), 10.97.21c *sárvāḥ saṃgátya vīrudho* (5th/8), and 10.130.7c *pū́rveṣām pánthām anudṛ́śya dhī́rā* (9th/11).

23 The sole example of an unexpected short form in the saṃhitāpāṭha is 6.4.5a *nítikti yó vāraṇám ánnam átti*. A few more short variants were later introduced by the padakāra (see Kuiper 1955: 4).

24 The rather mechanical character of these replacements is illustrated by 8.90.4b *vṛtrā́ bhū́ri nyṛ̥ñjáse* | scanned *bhū́rī ni-ṛ̥ñjáse*.

25 In this position there is a sharp contrast between e. g. the numerous attestations of *bhū́ri* (which is not attested with long final vowel at all) and the predominant (though not exclusive) use of *purú* (Kuiper 1955: 20–23).

26 Kuiper 1955: 19 cites 6.44.14a *asyá máde purú várpāṃsi vidvā́n* and 10.52.4b *ápamluktam bahú kṛ̥chrā́ cárantam* | with short plurals in the sixth syllable of triṣṭubhs. I have not found any other convincing cases where a singular reading seemed excluded.

to the gerunds, the short forms are preserved without exception at the end of all (not just the even) padas.

Although differing a lot in detail, all those forms from short vowel plus laryngeal thus have in common that they developed short variants at least in certain positions. With some of them (gerunds, *ī*-stem vocatives) these short variants even became the basic form, and the sole continuant in the later language. Why is it that only some forms have these variants, but others do not?[27] The proposed reconstruction of the singular nominatives of *ā*- and *ī*-stems does now allow for a straightforward answer: no short variants of these forms ever existed, because their vowels were always long, even when the laryngeal was dropped before a vowel or a pause. While clearly suggestive of the correctness of the proposed new reconstruction, this argument is nevertheless no proof against the traditional reconstruction, since there are also groups of forms with original short vowels plus laryngeal that have no short-vowel variants at all in the transmitted text.[28]

2.2 Metrical shortening of long vowels

While the preceding section dealt with a possible argument from the written text of the Rigveda for reconstructing a long vowel for *ā*- and *ī*-stem nominatives, this section is concerned with a possible counterargument from Rigvedic prosody. In Kuiper 1947: 208 not only the shortenings discussed above and the singular vocatives of *ī*-stems are mentioned as evidence for the proposed developments of word-final vowel plus laryngeal sequences, but also cases of "unexpected" short scansion of such vowels. He cites *prá devy ètu sūn̥ŕtā* (1.40.3b) as an example, but this is in fact not a very convincing case, since a long scansion would yield a structure that is actually more common.[29] The phenomenon is, however, still real, and better examples are attested.[30] But it is not clear whether such shortenings are just due to laryngeal sandhi before vowels, since examples such as *ná tásya rāyáḥ paryetā́sti* (7.40.3d, scanning *parieta asti*) occur too. Here it is the old lengthened

27 This question is already raised and left open by Kuiper 1955: 36: "In this connection we may stress the need of an examination of the sandhi of the feminine nominative endings *-ā* and *-ī*: why has the 'laryngeal-sandhi' thus far only been found in instrumental forms in *-tī* and *-(t)yā*?"

28 E. g. the neuter plural nominative-accusatives of *a*-stems and the paradigmatic instrumentals in *-ā*.

29 Van Nooten & Holland (1994: xvii) list 3641 instances of ⏑ – – – against 2786 instances of ⏑ – ⏑ – in the openings of eight-syllable padas.

30 E. g. *vadhū́r iyám pátim ichánty eti* (5.37.3a), where an *ī*-stem nominative is found in the ninth syllable of a triṣṭubh, a position where a long scansion is out of question.

grade of an *r*-stem singular nominative that fills the ninth syllable of the triṣṭubh cadence (see Oldenberg 1888: 465–469).

To establish whether long vowels of different sources show different distributions of this shortening rule, I collected the instances of the following long-vowel forms in the Rigveda. The first group comprises the forms already under discussion, the singular nominatives of *ā*- and *ī*-stems (< *-$\bar{V}H$ or *-$\check{V}H$); the second group comprises the neuter plural nominative-accusatives of *a*-, *i*-, and *u*-stems plus the singular instrumentals in *-(t)ǐ* (< *-$\check{V}H$); and the third group contains the singular nominatives of *ar*-, *an*-, and *in*-stems (< *-$\bar{V}$).[31] Table 2 shows the distribution of these forms at the end of padas, within padas before consonant, within padas contracted with a following vowel, and within padas in hiatus.[32]

Table 2. Attestations in pada final, pre-consonantic, and pre-vocalic positions

	nom. of *ī*-	**nom. of *ā*-**	**pl. n. of *i*-**	**pl. n. of *a*-**	**pl. n. of *u*-**	**instr. in *-ī/i***	**nom. of *in*-**	**nom. of *an*-**	**nom. of *ar*-**
final	219 (28%)	219 (19%)	15 (22%)	341 (19%)	65 (37%)	82 (42%)	49 (22%)	278 (27%)	151 (15%)
_C	455 (59%)	733 (65%)	46 (68%)	1226 (69%)	96 (55%)	85 (44%)	111 (50%)	647 (62%)	674 (69%)
contr. / _*V*	38 (5%)	138 (12%)	– (0%)	186 (10%)	– (0%)	3 (2%)	15 (7%)	94 (9%)	135 (14%)
hiatus / _*V*	64 (8%)	36 (3%)	7 (10%)	19 (1%)	13 (7%)	24 (12%)	46 (21%)	28 (3%)	17 (2%)
total	776	1126	68	1772	174	194	221	1047	977

31 The material was collected with the help of a lexical database of the Rigveda being developed at the University of Zurich based on Lubotsky 1997. The present paper would never have been written without him generously sharing his work, and without my colleagues in Zurich who have worked on the database over the past few years. They are planning to make this tool available to the scholarly community in the near future.

32 The figures include all attestations where I deemed the respective interpretation possible. Therefore some will probably be too high, especially those of the neuter plurals of *i*- and *u*-stems, because there are many situations where a decision between singular and plural is almost impossible. The attestations of *svastí* were left out of consideration because of the special problems they pose.

Contrasting the pada-final and pre-consonantal attestations, which are of no further interest for the present investigation, with all prevocalic examples reveals a significantly skewed distribution (Pearson's chi-square, $p < 0.001$),[33] but this is for the most part due to the very high number of *in*-stem nominatives before vowel, which are twice as frequent as expected.[34] Turning to just the prevocalic distribution, a rather clear (and unsurprising) general tendency emerges: contraction is much more common with *-ā* than with *-ī* and *-ū* ($p < 0.001$), which is probably simply a consequence of the frequencies of the word-initial vowels. Grouping the data according to the presence of a final laryngeal on the other hand reveals no significant correlation ($p > 0.1$).

Table 3 (p. 309) now focusses on the instances in hiatus. For each case I estimated the probability of its being long by the following, rather mechanical procedure: using the counts of metrical structures in van Nooten & Holland 1994: xvii–xviii, the percentage of long syllables in every position of lines of the common types was calculated,[35] and the data grouped according to those percentages. In Table 3 syllables with more than a 70% probability of being long are counted as "long", those under 30% as "short", and those in between as "undecided".[36]

The distribution seems certainly not random,[37] but the immediately visible peculiarity is again the aberrant distribution of the *in*-, and this time also the *an*-stems: short cases from *in*-stems are extremely common, but virtually nonexistent for *an*-stems. A similar but smaller difference exists also between the nominatives of *ī*- and *ā*-stems, and it is again the *-ī* that is shortened more often than

33 As with all subsequent distribution tests, this reads: "The odds that a random distribution (H_0) results in those figures are less than 1 in 1000."

34 Those 61 cases are also responsible for the marginal significance ($p < 0.05$) that results when **VH*- and **V̄*-forms are contrasted: 3582 **VH* /_#/*C*, 528 **VH* /_*V* opposed to 1910 **V̄* /_#/*C*, 335 **V̄* /_*V*. Grouping the data by vowel quality results in a not significant distribution, though the number of prevocalic attestations of *-ī/ū* is again above what is expected due to the *in*-stem forms. Their effect is cancelled to some extent by the relative rarity of *-ū* /_*V* in this case.

35 Those figures are: 8-syllable padas: 1. ≍, 2. 76%, 3. 66%, 4. 78%, 5. 6%, 6. 91%, 7. 6%, 8. ≍; openings of 11- and 12-syllable padas: 1. ≍, 2. 86%, 3. 49%, 4. 78%, 5. 56%; breaks of 11- and 12-syllable padas: 5. 30%, 6. 6%, 7. 73%, or 6. 15%, 7. 1%; cadences of 11- and 12-syllable padas: 8. 98%, 9. 1%, 10. 98%, 11. ≍/0%, 12. ≍. Special treatment was given to the unusual but consistent trochaic cadences in 8-syllable padas (– ⏑ – ≍), and dvipadāvirājs are scanned as ≍ – ⏑ – ≍.

36 This procedure is to some extent oversimplifying the matter, since it misses some known dependencies between syllables beyond the general patterns, but it will do as a first approximation and allow us to grasp the general picture.

37 The probability of the overall table is hard to assess exactly though, because a Fisher's Exact test runs into memory limitations. The chi-square test results in $p < 0.001$ ($p = 6.892 \cdot 10^{-8}$), but there are too many cells with low numbers to allow for a really reliable result.

Table 3. Length in hiatus

	nom. of *ī-*	nom. of *ā-*	pl. n. of *i-*	pl. n. of *a-*	pl. n. of *u-*	instr. in *-ī/i*	nom. of *in-*	nom. of *an-*	nom. of *ar-*
long	17 (27%)	21 (58%)	– (0%)	8 (42%)	5 (38%)	7 (29%)	8 (17%)	22 (79%)	8 (47%)
undec.	26 (41%)	9 (25%)	3 (43%)	7 (37%)	2 (15%)	7 (29%)	7 (15%)	5 (18%)	4 (24%)
short	21 (33%)	6 (17%)	4 (57%)	4 (21%)	6 (46%)	10 (42%)	31 (67%)	1 (4%)	5 (29%)
average	.47	.71	.21	.60	.46	.44	.25	.88	.59
total	64	36	7	19	13	24	46	28	17

the *-ā*. This effect proves highly significant ($p < 0.001$) when the data is collapsed to contrast the vowel qualities (Table 4).[38] The frequency of prevocalic shortening is thus clearly correlated with vowel quality (Cramer's $V = 0.5494$).

Table 4. Length in hiatus: *-ā* vs.*-ī* vs. *-ū*

	-ā	*-ī*	*-ū*
long	59	32	5
undec.	25	43	2
short	16	66	6
total	100	141	13

Grouping the data by category on the other hand reveals no conclusive tendencies.[39] The only grouping that at first sight produces an at least marginally significant result ($p < 0.05$) is the one comparing all forms in *-*H* to those in *-$\bar{V}$. But this 'correlation' is not what could have been expected, because when we compare the

38 The forms in *-ū* are distributed roughly as expected, but the numbers are too small to be really conclusive. It is therefore not surprising that contrasting *-ā* vs. *-ī* vs. *-ū*, or *-ā* vs.*-ī* and *-ū*, or *-ā* vs. just *-ī* all give very similar results.

39 The following groupings were tested: nom. of $\bar{V}$- vs. n. pl. vs. instr. vs. nom. of *C*-, nom. of $\bar{V}$- vs. n. pl. vs. nom. of *C*-, nom. of $\bar{V}$- vs. n. pl. and instr. vs. nom. of *C*-, all with Fisher's Exact test $p > 0.1$.

figures with expected results, the cell contributing most to the actual result is the low number of undecided cases from *-$\bar{V}$ (see Table 5). When combining "short" and "undecided" as "not long", or "long" and "undecided" as "not short", the significance disappears ($p > 0.1$ in both cases). Finally I also compared the figures of the groups in *-*H* alone; again there is no significant correlation. If anything, the neuter plurals and instrumentals are shortened slightly more often than expected, but this can easily be due to chance.

Table 5. Length in hiatus: *-*H* vs. *-$\bar{V}$

	-H			***-$\bar{V}$***		
long	58		58	38		38
not short		112			54	
undec.	54			16		
not long			105			53
short	51	51		37	37	
total		163			91	

The overall conclusion is very clear. There was a rule of vowel shortening before another vowel that affected different forms to different extents, but without regard to their respective origins from old long vowels or from sequences of vowel plus laryngeal. The attested metrical shortenings of *ā*- and *ī*-stem nominatives are thus no argument against a reconstruction of these forms with an original long vowel. What needs to be investigated further is the exact distribution of this prosodic option, and especially how the very special situation with *in*- and *an*-stems can be explained. Why is the nominative of *in*- so common before vowels, and why is it shortened so much more often than the other groups? Why is the nominative of *an*-stems almost never shortened? And is the higher frequency of shortening of *-ī* just a consequence of the higher frequency in hiatus, or rather an independent effect? Do the distributions differ between the different books of the Rigveda? The study of formulaic expressions will certainly play a major role in this task. For instance, looking more closely at the 22 cases of long *an*-stem nominatives we find 5 × *vr̥trahā́*, 2 × *rā́jā*, 1 × *yúvā*, and 1 × *brahmā́* before a caesura in the fourth syllable with a form of *índra*- filling the break.[40] But this approach alone will not suffice for explaining the wealth of attestations of shortened *in*-stem nom-

40 1.186.6c, 2.20.7a, 8.96.20a, 8.96.21a, 10.74.6b; 1.174.1a, 1.178.2a; 2.20.3a; 8.16.7a.

inatives, especially in cadences of 11- and 12-syllable padas,[41] because, while *vājī́* (15×), *dyumnī́* (5×), and *śuṣmī́* (4×) clearly dominate the picture,[42] the formulaic elements involved are all much more varied and flexible.

Acknowledgement: My thanks go to the anonymous reviewers, from whose comments this paper has benefitted both in content and in style. Of course, the usual disclaimers apply.

Bibliography

Beekes, Robert S. P. (1985). *The Origins of the Indo-European Nominal Inflection*. Innsbruck: Institut für Sprachwissenschaft.

Fritz, Matthias (1998). "Die urindogermanischen *s*-Stämme und die Genese des dritten Genus". In: *Sprache und Kultur der Indogermanen. X. Fachtagung der Indogermanischen Gesellschaft, Innsbruck, 22.–28. September 1996*. Innsbruck: Institut für Sprachwissenschaft, 255–264.

Harðarson, Jón Axel (1987). "Zum urindogermanischen Kollektivum". In: *Münchener Studien zur Sprachwissenschaft* 48, 71–113.

Hock, Wolfgang (2004). "Baltoslavisch. I. Teil: Phonologie". In: *Kratylos* 49, 1–32.

Kim, Ronald I. (2014). "A tale of two Suffixes: **"h_2"*, **"ih_2"*, and the evolution of feminine gender in Indo-European". In: *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*. Ed. by Sergio Neri & Roland Schuhmann. Leiden: Brill, 115–136.

Klingenschmitt, Gert (1992). "Die lateinische Nominalflexion". In: *Latein und Indogermanisch*. Akten des Kolloqiums der Indogermanischen Gesellschaft, Salzburg, 23.–26. September 1986. Ed. by Oswald Panagl & Thomas Krisch. Innsbruck: Institut für Sprachwissenschaft, 89–135.

Kortlandt, Frederik H. H. (1985). "Long vowels in Balto-Slavic". In: *Baltistica* 21.2, 112–124.

Kortlandt, Frederik H. H. (1997). "PIE lengthened grade in Balto-Slavic". In: *Festschrift for Eric P. Hamp*. Ed. by Douglas Q. Adams. Vol. 2. Washington D. C.: Institute for the Study of Man, 26–31.

Kuiper, Franciscus B. J. (1947). "Traces of laryngeals in Vedic Sanskrit". In: *India Antiqua*. A volume of oriental studies presented by his friends and pupils to Jean Philippe Vogel on the occasion of the fiftieth anniversary of his doctorate. Ed. by F. D. K. Bosch. Leiden: Brill, 198–212.

Kuiper, Franciscus B. J. (1955). "Shortening of final vowels in the Rigveda". In: *Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afd. Letterkunde* 18.11, 1–37 = 253–289.

41 19 attestations, besides 11× in the fifth of 8-syllable padas and once in the third of a dvipadāvirāj.

42 1.74.8a, 1.133.7e, 1.163.12a, 4.38.10c, 6.7.3a, 7.44.4a, 7.90.2d, 8.32.18b, 9.64.29b, 9.86.11a, 9.87.4d, 9.97.45b, 9.107.5c, 9.109.17a, 10.34.4b; 1.36.8c, 8.89.2b, 8.103.9b, 10.159.4b, 10.174.4b; 4.22.1b, 4.22.4c, 7.56.24a, 9.71.1a.

Ledo-Lemos, Francisco José (2000). *Femininum Genus. A Study of the Origins of the Indo-European Feminine Grammatical Gender*. München: Lincom.

Litscher, Roland (2009). "Die Genese des dritten Genus: ein neuer Versuch". In: *Protolanguage and Prehistory*. Akten der XII. Fachtagung der Indogermanischen Gesellschaft, vom 11. bis 15. Oktober 2004 in Krakau. Ed. by Rosemarie Lühr & Sabine Ziegler. Wiesbaden: Reichert, 271–285.

Litscher, Roland (2014). "Voraussetzungen für ein feminines Genus und Implikationen für das Kategoriensystem des frühindogermanischen Nomens". In: *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*. Ed. by Sergio Neri & Roland Schuhmann. Leiden: Brill, 137–165.

Lubotsky, Alexander M. (1997). *A Ṛgvedic Word Concordance*. 2 vols. New Haven: American Oriental Society.

Luraghi, Silvia (2009). "Indo-European nominal classification. From abstract to feminine". In: *Proceedings of the 20th Annual UCLA Indo-European Conference*. Ed. by Stephanie W. Jamison, H. Craig Melchert & Brent Vine. Bremen: Hempen, 115–131.

Luraghi, Silvia (2014). "Gender and word formation. The PIE gender system in cross-linguistic perspective". In: *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*. Ed. by Sergio Neri & Roland Schuhmann. Leiden: Brill, 199–232.

Mayrhofer, Manfred (1986). *Indogermanische Grammatik*. Vol. 1.2: *Lautlehre [Segmentale Phonologie des Indogermanischen]*. Heidelberg: Winter.

Melchert, H. Craig (2014). "PIE *-*eh*$_2$ as an "individualizing" Suffix and the Feminine Gender". In: *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*. Ed. by Sergio Neri & Roland Schuhmann. Leiden: Brill, 257–272.

Neri, Sergio & Roland Schuhmann, eds. (2014). *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*. Leiden: Brill.

Van Nooten, Barend A. & Gary B. Holland (1994). *Rig Veda. Metrically Restored Text with an Introduction and Notes*. Cambridge, MA: Harvard University Press.

Nussbaum, Alan J. (2014). "Abstract, collective, neuter plural. Some remarks on each (Expanded Handout)". In: *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*. Ed. by Sergio Neri & Roland Schuhmann. Leiden: Brill, 273–306.

Oldenberg, Hermann (1888). *Die Hymnen des Rigveda*. Vol. 1: *Metrische und textgeschichtliche Prolegomena*. Berlin: Hertz.

Tichy, Eva (1993). „Kollektiva, Genus femininum und relative Chronologie im Indogermanischen". In: *Historische Sprachforschung* 106, 1–19.

Widmer, Paul (2005). "Der altindische *vṛkí*-Typus und hethitisch *nakkǐ*-: Der indogermanische Instrumental zwischen Syntax und Morphologie". In: *Die Sprache* 45, 190–208.

Zubatý, Josef (1889). "Der Quantitätswechsel im Auslaute vedischer Wörter (Fortsetzung)". In: *Wiener Zeitschrift für die Kunde des Morgenlandes* 3, 86–92, 151–162, 281–312.